



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 30, 2017 – 06:11 PM EDT

PDB ID : 3MDZ
Title : Crystal Structure of Human Carbonic Anhydrase VII [isoform 1], CA7
Authors : Ugochukwu, E.; Shafqat, N.; Pilka, E.; Chaikuad, A.; Krojer, T.; Muniz, J.; Kim, J.; Bray, J.; Bountra, C.; Arrowsmith, C.H.; Weigelt, J.; Edwards, A.; von Delft, F.; Carpenter, E.P.; Yue, W.W.; Oppermann, U.; Structural Genomics Consortium (SGC)
Deposited on : unknown
Resolution : 2.32 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : **FAILED**
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030345

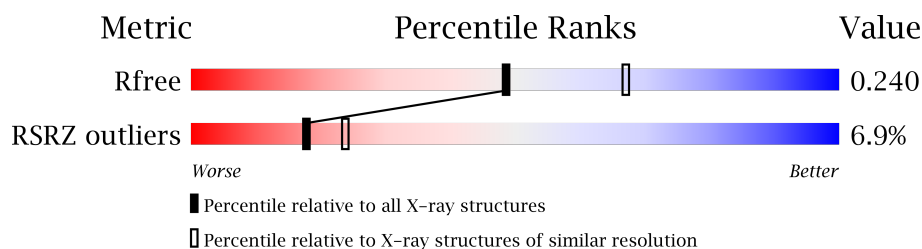
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.32 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	4787 (2.34-2.30)
RSRZ outliers	101464	4814 (2.34-2.30)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GOL	A	266	-	-	-	X
4	GOL	A	267	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2077 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Carbonic anhydrase 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	259	Total	C	N	O	S	0	1	0
			1994	1278	349	358	9			

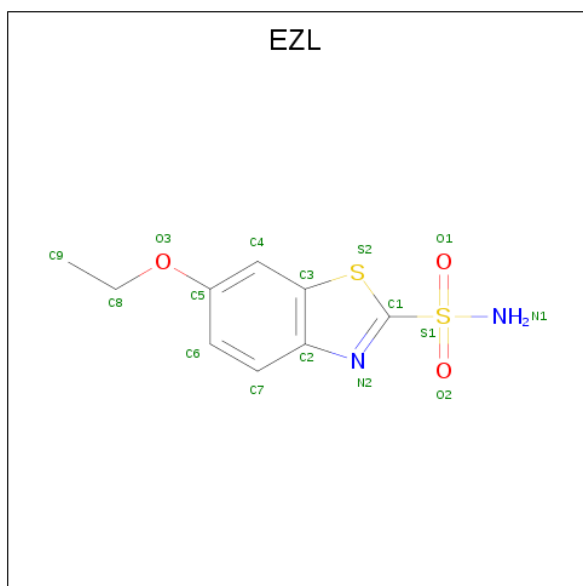
There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-18	MET	-	EXPRESSION TAG	UNP P43166
A	-17	HIS	-	EXPRESSION TAG	UNP P43166
A	-16	HIS	-	EXPRESSION TAG	UNP P43166
A	-15	HIS	-	EXPRESSION TAG	UNP P43166
A	-14	HIS	-	EXPRESSION TAG	UNP P43166
A	-13	HIS	-	EXPRESSION TAG	UNP P43166
A	-12	HIS	-	EXPRESSION TAG	UNP P43166
A	-11	SER	-	EXPRESSION TAG	UNP P43166
A	-10	SER	-	EXPRESSION TAG	UNP P43166
A	-9	GLY	-	EXPRESSION TAG	UNP P43166
A	-8	VAL	-	EXPRESSION TAG	UNP P43166
A	-7	ASP	-	EXPRESSION TAG	UNP P43166
A	-6	LEU	-	EXPRESSION TAG	UNP P43166
A	-5	GLY	-	EXPRESSION TAG	UNP P43166
A	-4	THR	-	EXPRESSION TAG	UNP P43166
A	-3	GLU	-	EXPRESSION TAG	UNP P43166
A	-2	ASN	-	EXPRESSION TAG	UNP P43166
A	-1	LEU	-	EXPRESSION TAG	UNP P43166
A	0	TYR	-	EXPRESSION TAG	UNP P43166
A	1	PHE	-	EXPRESSION TAG	UNP P43166
A	2	GLN	-	EXPRESSION TAG	UNP P43166
A	3	SER	-	EXPRESSION TAG	UNP P43166
A	4	MET	-	EXPRESSION TAG	UNP P43166

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

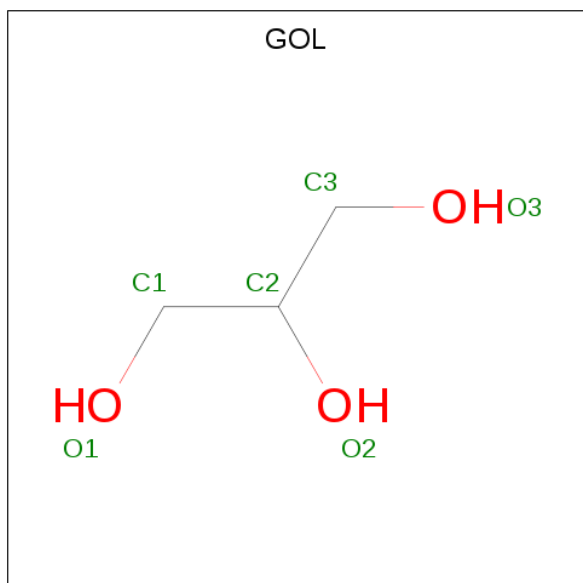
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is 6-ethoxy-1,3-benzothiazole-2-sulfonamide (three-letter code: EZL) (formula: $C_9H_{10}N_2O_3S_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			16	9	2	3	2		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	48	Total	O	0	0
			48	48		

MolProbity failed to run properly - this section is therefore empty.

3 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	69.71Å 97.75Å 82.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.74 – 2.32 19.74 – 2.32	Depositor EDS
% Data completeness (in resolution range)	99.2 (19.74-2.32) 99.4 (19.74-2.32)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.12 (at 2.33Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.200 , 0.252 0.195 , 0.240	Depositor DCC
R_{free} test set	588 reflections (4.97%)	DCC
Wilson B-factor (Å ²)	48.5	Xtriage
Anisotropy	0.105	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 46.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	2077	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.14% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

4.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

4.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond

length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	EZL	A	264	2	14,17,17	1.46	2 (14%)	16,25,25	2.55	7 (43%)
4	GOL	A	265	-	5,5,5	0.23	0	5,5,5	0.73	0
4	GOL	A	266	-	5,5,5	0.65	0	5,5,5	0.46	0
4	GOL	A	267	-	5,5,5	0.54	0	5,5,5	0.76	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EZL	A	264	2	-	0/3/9/9	0/2/2/2
4	GOL	A	265	-	-	0/4/4/4	0/0/0/0
4	GOL	A	266	-	-	0/4/4/4	0/0/0/0
4	GOL	A	267	-	-	0/4/4/4	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	264	EZL	C1-S1	-3.44	1.74	1.78
3	A	264	EZL	O2-S1	2.36	1.48	1.43

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	264	EZL	O2-S1-O1	-5.06	109.42	118.70
3	A	264	EZL	O3-C5-C4	-2.70	115.06	123.99
3	A	264	EZL	C3-C2-N2	2.22	113.16	108.04
3	A	264	EZL	C4-C3-S2	2.56	129.55	124.84
3	A	264	EZL	C1-S1-N1	3.47	114.87	108.23

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data [i](#)

5.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	259/281 (92%)	0.25	18 (6%) 18 24	26, 52, 88, 107	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	75	SER	4.8
1	A	11	GLN	4.5
1	A	5	HIS	3.6
1	A	12	ASP	3.2
1	A	189	ALA	3.1

5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	GOL	A	267	6/6	0.75	0.20	3.52	61,63,64,64	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	GOL	A	266	6/6	0.82	0.28	2.54	20,25,26,28	6
4	GOL	A	265	6/6	0.80	0.17	1.00	56,61,62,65	0
3	EZL	A	264	16/16	0.97	0.12	-0.13	38,44,47,48	0
2	ZN	A	263	1/1	0.99	0.06	-	35,35,35,35	0

5.5 Other polymers [i](#)

There are no such residues in this entry.